CRM D0005586.A2/Final
February 2002

An Examination of the Ship Material Condition Metrics (SMCM) Methodology

Peter J. Francis • James M. Jondrow
William H. Sims • John D. Keenan
Geoffrey B. Shaw

CNA
4825 Mark Center Drive • Alexandria, Virginia 22311-1850

Approved for distribution: February 2002

Alan J. Marcus, Director
Infrastructure and Readiness Team
Resource Analysis Division

This document represents the best opinion of CNA at the time of issue.
It does not necessarily represent the opinion of the Department of the Navy.

Approved for Public Release; Distribution Unlimited. Specific authority: N00014-00-D-0700.
For copies of this document call: CNA Document Control and Distribution Section at 703-824-2123.

Copyright 2002 The CNA Corporation

Contents

Summary

There is a perceived need for a single metric that represents the operational mechanical and electrical (M&E) readiness of ships. Such a number could be useful in maintenance planning, programming, and execution; in evaluating whether the fleet is ready for a contingency; and in spotting systematic readiness deficiencies and making associated resource decisions. One effort to develop such a metric is the Ship Material Condition Metrics (SMCM) initiative developed at the Naval Warfare Assessment Station (NWAS) at Corona, California. NWAS was assisted in this effort by the Fleet Technical Support Center, Pacific (FTSCPAC). As a pilot project, the formulas for the metric have been applied to the evaluation of the USS *Lake Champlain* (CG-57).[1] OpNav N81 asked CNA to evaluate the way the metric was constructed.

The NWAS model aims to roll up “readiness” evaluations of small pieces of equipment into a single number for the entire ship M&E. These evaluations are obtained by an inspection team; these inspections are intended to occur once during a ship’s interdeployment training cycle (IDTC). The formula is based on one used in the Troubled Systems Process (TSP) which NWAS developed for combat systems.

We assessed the model on two levels. In the first, we evaluated the NWAS model parameters and model structure for a significant subset of the ship systems. A small panel of naval engineers from the CNA Naval Studies Group assisted us in our investigations. As a result of our investigations, we can suggest modifications to the formulas which should permit better handling of redundant systems and should better represent the criticality of particular subsystems. This

1. See *Assessment of Equipment Condition (AEC) Troubled Systems Process (TSP) Report for USS Lake Champlain (CG-57) Pilot*, Naval Warfare Assessment Station, 22 February 1999.

part of the evaluation took the basic structure and goals of the model as given and represents an examination of the model "on its own terms."

In the other part of our evaluation, we took a step back and critically reviewed the fundamental structure of the calculations, formulas, and numerical scales. Our aim was to evaluate whether the SMCM methodology would give measures that are meaningful, consistent, and useful. We considered the way in which scales, formulas, and parameters were developed; the way in which data are collected; and how these data might best be used. We also considered whether the need for material condition metrics is best served by generating a single number, and what that number should represent.

Results

We evaluated the model on its own terms and learned the following:

- The NWAS roll-up equation, used with the NWAS model parameters, does not yield realistic roll-up results. This appears to be a consequence of the formula's parameters (criticality and weights), which appear to be unrealistic for many items. When the NWAS formula is used with the CNA panel's model parameters and our proposed modifications to treat redundant subsystems, it does yield reasonably realistic results.
- The equipment operational capability (EOC) metric is probably satisfactory for measuring current operating capability but may not be adequate for assessing equipment condition.
- The reliability of EOC codes assigned by technicians to low-level equipment should be evaluated. This should improve the credibility of the process.

We evaluated the fundamental structure of the model and obtained these results:

- The functional forms and numerical scales used in the SMCM (and TSP) methodology appear to be arbitrary constructs with little or no analytical underpinnings.

- Using a simple example, we demonstrated that different EOC scales can produce different rank orderings of material conditions even when other model parameters are specified correctly. The implication of this result is that an arbitrarily chosen numerical scale or roll-up function may produce incorrect results even after other parameters are correctly specified.

- It may be possible to develop some of the missing analytical bases for the SMCM model structure by using what we have come to call the "truth table" methodology. Truth tables are just descriptions of the overall condition of a ship and the condition of its systems and subsystems. Once we find reliable descriptions of these conditions, they can be used to evaluate the extent to which numbers and formulas can meaningfully replicate the descriptions.

- The more disparate the components of any composite measure or index, the more difficult it is to find a meaningful way to combine those components into a single number, and the more difficult it is to interpret the resulting metric. In particular, it appears that the information captured in the system reflects a variety of dimensions: the extent to which the system is operational, the impact on higher level systems, problems such as improper installation or labeling, and who should fix the system. These multiple dimensions may naturally lead to multiple indexes.

- The current methodology, which is based on an inspection by a team, results in very infrequent evaluations of each ship and largely fails to take advantage of modern information technology, which should be capable of producing ship material condition reports almost in real time.

Recommendations

Our recommendations are as follows:

- Consider modifying the formula for the calculations, especially the handling of redundancy, and the parameters such as

weights and criticality to account for the issues uncovered by our panels.

- Ensure that the technicians who assign low-level EOC codes are thoroughly trained in the evaluation methodology.
- Ensure that definitions of the EOC metric are consistent with policy objectives and are clearly stated.
- Ensure that equipment is not marked down on EOC for superficial reasons such as improper labeling or failing to install according to instructions. This was common in the prototype.
- Improve the credibility of the model by empirically measuring the reliability (consistency in repeated evaluations) of the EOC codes assigned by trained technicians using the method of repeated evaluations by different individuals.
- Apply the "truth table" methodology to the revised index, to make sure that it captures common sense descriptions of ship condition.
- Consider the likelihood that the needs of the Navy will be better served by treating material condition as a multidimensional characteristic and reporting more than one material condition measure for a ship. However, we believe that this question deserves further investigation and so will not make a definite recommendation on dimensionality at this time. NWAS is discussing a proposal to develop material condition metrics by mission area; depending on how it is implemented, this proposal could help settle the dimensionality question.
- We recommend that the SMCM initiative be developed in close conjunction with the implementation of shipboard Enterprise Resource Planning (ERP) installations, since much of the information necessary to evaluate the material condition of a ship should be available in real time through the ERP systems. The system would be much more useful if the data were available continually.

Background

The NWAS model aims to roll up ship M&E "readiness" evaluations of small pieces of equipment into a single number for the entire ship. The model is based on one used in the Troubled Systems Process (TSP) that NWAS developed for use with combat systems.

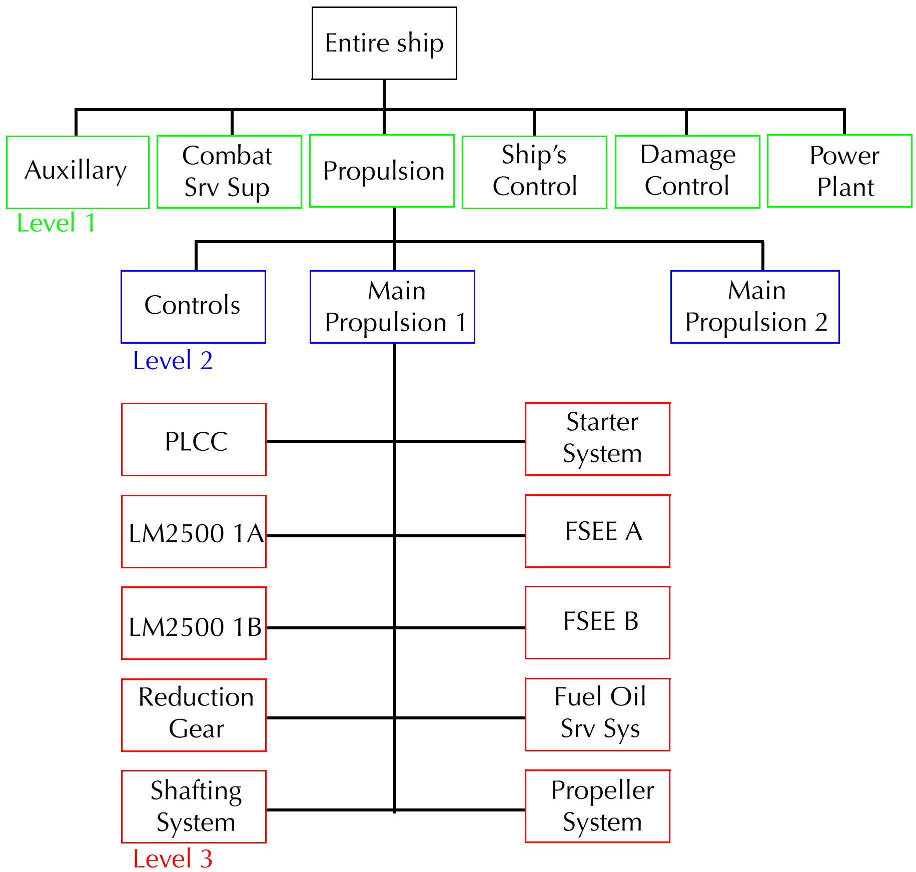
Each piece of equipment is assigned a level in the structural tree. The structural tree consists of five levels and is made up of "parents" and "children." Parents may be thought of as systems and children as subsystems. Parents at one level may be children in a higher level system. FTSCPAC technicians perform the evaluations on the condition of individual pieces of equipment. The condition of each piece is assessed at the lowest level in the structural tree on an EOC metric. The EOC numbers are then rolled up to get estimates for each level and for the entire ship using an equation and model parameters. The model parameters are the weights and criticality factors associated with each piece of equipment at each level of the structural tree.

Figure 1 shows what a structural tree is like. Auxillaries, combat service support, propulsion, ship's control, damage control, and power plant are the systems that make up level 1. NWAS considers each of these elements to be a "parent."

If we decompose the propulsion system, we see that, at level 2, it is comprised of controls, main propulsion 1, and main propulsion 2. NWAS considers each of these elements to be “parents.”

If we decompose “main propulsion 1,” we find that it is comprised of PLCC, LM 2500 1A, LM 2500 1B, reduction gear, shafting system, starter system, FSEE A, FSEEB, fuel oil service system, and propeller system. In this example, main propulsion 1 would be a "parent," and reduction gear would be a "child.” (Figure 1 does not show the other branches of level 1 and all of the level 4 and level 5 subsystems.)

Figure 1. Structural tree (level 1 and selected elements of levels 2 and 3)



Definition of the EOC scale

Each element of the tree at the lowest level is assigned an EOC code. The EOC metric is defined as shown in table 1.

Note that the EOC metric is designed to focus on the current operating capability of the equipment—not on the material condition of the equipment. This may appear to be a philosophical point, but it is an important distinction. For example, the operating capability of the equipment may be fine on the date of inspection although the material condition may be such that breakage is likely to occur in the near future.

Table 1. Definitions of EOC metric

EOC range	Condition	Detail
$EOC \geq 0.8$	Operable	Item functions within parameters of PMS and/or passes all operational tests
$0.6 \leq EOC < 0.8$	Minor problem	Item slightly below PMS and/or unable to obtain optimum operational standards
$0.4 \leq EOC < 0.6$	Limited capability	Item fails PMS and/or operational tests and has significantly reduced output or restricted availability
$0.2 \leq EOC < 0.4$	Major problem	Item will not operate; may provide limited output if manually operated
$EOC < 0.2$	Inoperative	

It appears that the EOC codes as currently assigned do not fully reflect the material condition of the ship. Some may assume that because the EOC codes reflect condition within PMS that material condition is automatically incorporated in the EOC ratings. This may well be the case for combat systems; however, it appears not to be the case for M&E items. We understand that PMS for M&E items generally does not address material condition. It is thought that INSURV is attempting to incorporate material condition into PMS for M&E systems, but this change has not yet been accomplished.

Policy-makers should be clear on just what they want these EOC codes to measure. If the desire is to measure a snapshot of current operating capability, then the current definitions may be adequate. If the desire is to measure material condition, then the current definitions may not be adequate.

Definition of a roll-up equation

The NWAS roll-up equation is different depending on whether or not the family contains a critical item. The definitions are given in table 2.

Table 2. NWAS roll-up equation

Critical items in family	Roll-up equation[a]
No	Roll-up $EOC = \Sigma\, EOC_i\, Weight_i \,/\, \Sigma\, Weight_i$
Yes	Roll-up $EOC = EOC_{lc}\, [Weight_{lc} + \Sigma\, EOC_j\, Weight_j] \,/\, \Sigma\, Weight_i$

a. Where EOC_{lc} is the EOC of the critical item with the lowest EOC rating, the "j" represents all other items, and the "i" represents all items. Note that the roll-up EOC can never be greater than the EOC of the lowest critical item.

Definitions of model parameters

Table 3 defines the model parameters. NWAS and FTSCPAC developed these parameters with the help of technical experts.

Table 3. Model parameters

Factor	Definition
Criticality	An item is critical if the parent cannot operate without the child. The item is non-critical if the parent can operate without the child, perhaps with a reduced capability
Weight	The weight factor ranges from 1 to 99 and reflects the relative importance of the child to other children in the family

Assessing the model on its own terms

Potential inaccuracies could result from problems with the structure of the roll-up equation, the model parameters (criticalities and weights), or the EOCs assigned at the lowest level by FTSCPAC technicians. We will examine each of these issues in turn.

Model parameters

In this section, we describe our analysis of the model parameters. Our analysis is based on an independent assessment of the criticality and weight factors of all level-1 and level-2 items, as well as selected items from level 3. Our study team made the assessment with the assistance of a panel of naval engineers from the CNA Naval Studies Group.

We examined in detail the roll-ups for the systems listed in table 4.

Table 4. Levels, systems, and subsystems examined

Level	System/subsystem	Number of subsystems
1	Entire ship	6
2	Propulsion	3
3	Main propulsion 1	10
2	Auxillaries	11
2	Combat systems support	8
2	Ship's control	2
2	Power plant	5
2	Damage control	2
3	Firemain	9
Total		56

Level 1 roll-up (entire ship) consists of six subsystems: propulsion, auxillaries, etc. Each of these six subsystems is a level-2 system and

may contain other subsystems. For example, the propulsion subsystem (level 2) consists of three subsystems, only one of which, main propulsion 1, is examined. In total, we examined the roll-up EOCs for 56 systems or subsystems. The list includes all of the level-1 and level-2 systems, as well as selected level-3 systems.

Methodology

We evaluated each of the 56 subsystems listed in table 4 in terms of our judgement as to the reasonableness of the model parameters and the roll-up produced by these parameters. We based our judgements on our experience and on the input of our panel of naval engineers.

We will illustrate our methodology by applying it to the level-1 roll-up. Table 5 lists the level-1 model parameters.

Table 5. Level-1 model parameters

	Criticality		Weight	
System	NWAS	CNA panel	NWAS	CNA panel
Auxillaries	N	N	20	1
Combat service support	N	C	20	10
Ship's control	N	C	20	19
Damage control	N	C	20	5
Propulsion	N	C	20	20
Power plant	N	C	20	45

The first column lists the systems. Each system will be designated as critical (C), or non-critical (N). The second column lists the criticality as specified in the NWAS model. The third column lists the judgement of our panel of naval engineers. For example, the NWAS model considers the propulsion system non-critical to the overall EOC code of the ship. Our panel took strong exception to that designation. They noted that the Required Operational Capability (ROC) of the ship involves steaming at designated speeds and listed propulsion as a critical element. Without propulsion, the ship has no meaningful operational capability. For similar reasons, we also designated combat service support, ship's control, damage control, and power plant as

critical elements in rolling-up an overall EOC code for the ship's M&E status. In our judgement, the loss of any of these systems renders the ship non-operational, and this reality should be reflected in the model roll-up of an overall EOC code.

Designation of an item as critical has a major impact on the roll-up model because in the NWAS equation, the roll-up EOC code can never be greater than the EOC code of the lowest rated critical item. In table 6, we summarize the criticality designations for the 56 systems and subsystems that we examined.

Table 6. Summary of criticality designations

Level	System	Number of subsystems	Number of subsystems designated as cricital	
			NWAS	CNA
1	Entire ship	6	0	5
2	Auxillaries	11	0	3
2	Combat service support	8	0	6
2	Ship's control	2	0	1
2	Damage control	2	0	1
2	Propulsion	3	0	1 (2)[a]
2	Power plant	5	0	1
3	Firemain	9	0	0 (1)[a]
3	Main propulsion 1	10	6	6 (7)[a]
	Totals	56	6	24 (27)

a. The number of critical systems if certain redundant systems are designated as critical.

As we see from table 6, none of the level-1 or level-2 systems or subsystems are designated as critical in the NWAS model. We, on the other hand, consider many level-1 and level-2 systems to be critical. As previously noted, this difference will have a major effect on the model results. We also believe that the model could be improved if certain redundant systems were classified as critical or conditionally critical.

We also see from table 6 that the panel arrived at a different set of weights for many of the subsystems. Although the model ultimately normalizes any set of weights, we found it to be a useful discipline to constrain the weights to sum to 100 for each set of subsystems. We will see in later discussion that the impact of differing sets of weights is rather small.

Effect of model parameters on the roll-up equation

In this section, we examine the impact of differing sets of criticalities and weights on the results from the NWAS roll-up equation. We assigned a set of hypothetical EOCs to each subsystem at various levels and asked our panel of experienced naval engineers to estimate a rolled-up EOC code for that set of subsystem EOC codes. We then calculated the roll-up using the NWAS equation and various combinations of NWAS and panel criticalities and weights.

Clearly, there are considerable uncertainties in any estimated roll-up number arrived at by our panel. Nonetheless, the panel deliberated at considerable length to arrive at an informed consensus, and we consider their estimates to have validity. In any event, if the model is to be accepted by the larger community, it must produce estimates that seem reasonable to persons such as our panel of naval engineers.

We chose the hypothetical EOC codes in the NWAS metric to be:

1.0(Operable)

0.5(Limited capability)

0.0 (Inoperative)

We illustrate the procedure using the input for the level-1 roll-up for the entire ship. The panel was given the list of hypothetical EOC codes for all level-1 subsystems as shown in table 7.

For example, under hypothetical EOC condition "A," we assume that all subsystems are in operable condition except for propulsion, which we designate as 0.5, i.e., "limited capability."

Table 7. Hypothetical subsystem EOC codes for level-1 roll-up (entire ship)

Subsystem	Hypothetical EOC conditions A	B	C	D	E	F
Auxillaries	1.0	1.0	1.0	1.0	1.0	0.5
Combat systems support	1.0	1.0	1.0	1.0	0.5	0.5
Ship's control	1.0	1.0	1.0	0.5	0.5	0.5
Damage control	1.0	1.0	1.0	1.0	1.0	0.5
Propulsion	0.5	0.0	0.5	0.5	0.5	0.5
Power plant	1.0	1.0	0.5	0.5	0.5	0.5

The panel was asked to estimate the overall M&E EOC code for the entire ship based on the NWAS definitions and the hypothetical EOC codes as shown in table 7 for each of the conditions A through F.

We also used the NWAS equation and the level-1 model parameters shown in table 8 to estimate the roll-up EOC code for the entire ship. We compare the results of these two methods of estimating the roll-up EOC code for the entire ship in table7.

Table 8. Comparison of level-1 roll-up EOC codes from NWAS equation with panel estimates

Parameters		Calculated roll-up for hypothetical EOC condition					
Criticality	Weights	A	B	C	D	E	F
NWAS	NWAS	.92	.83	.83	.75	.67	.50
NWAS	Panel	.90	.80	.68	.58	.53	.50
Panel	Panel	.50	.00	.45	.40	.38	.36
Panel estimate of roll-up		.67	.00	.67	.50	.25	.10

For example, we see that for EOC condition "A" in table 7, the NWAS equation using NWAS criticalities and NWAS weights estimates a roll-up EOC code of 0.92 for the entire ship. In comparison, using the same NWAS equation but with the panel's criticalities and the panel's

weights, we calculate a roll-up EOC code of 0.50. These two estimates can be compared to the panel's estimate of the roll-up of 0.67.

It is instructive to compare the roll-ups shown in the first and second rows of data. They differ only in that NWAS weights were used in the first row and the panel's weights were used in the second row. These data suggest that the results are not very sensitive to different weights.

The overall agreement of roll-ups using the NWAS parameters with direct estimates by the panel is not good. In contrast, using the panel's parameters (in the NWAS equation) seems to lead to reasonably good agreement with the panel's overall estimates. It appears that the lack of agreement when using the NWAS parameters is due to the criticality factors and not to the weight factors. Generally, the NWAS parameters lead to overestimation of the roll-up EOC codes. This appears to be the result of the NWAS designation of all major subsystems as non-critical to the overall M&E EOC of the ship (see table 5).

We carried out a similar analysis for all of the systems and subsystems shown in table 4. This gave us 51 estimated roll-ups using the NWAS parameters and 51 using the panel's parameters. The NWAS equation was used for all roll-ups. For each roll-up, we computed the difference between the roll-up EOC code and the roll-up that had been

directly estimated by the panel. We show the distribution in these differences in figures 2 and 3.

Figure 2. Difference in EOC roll-up and panel estimate: NWAS criticalities and NWAS weights

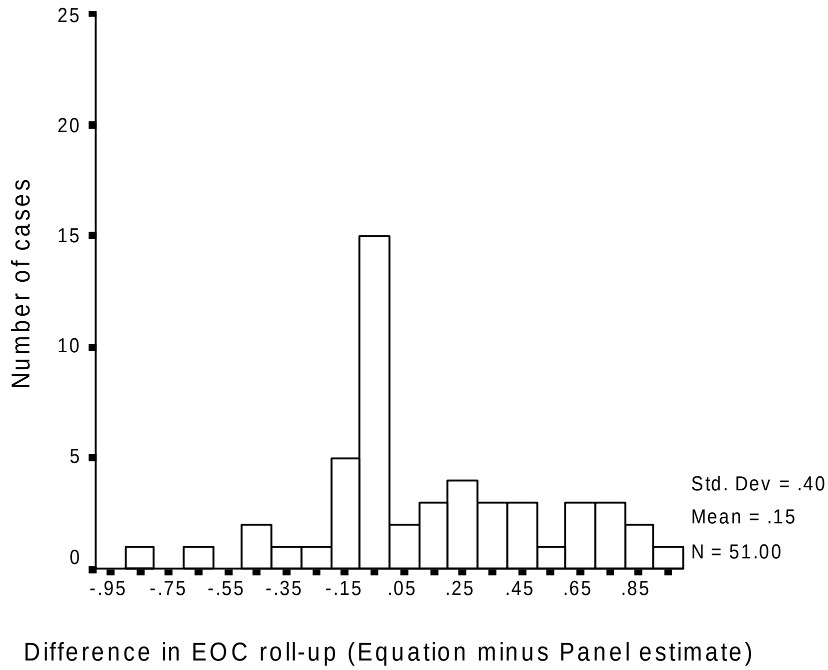
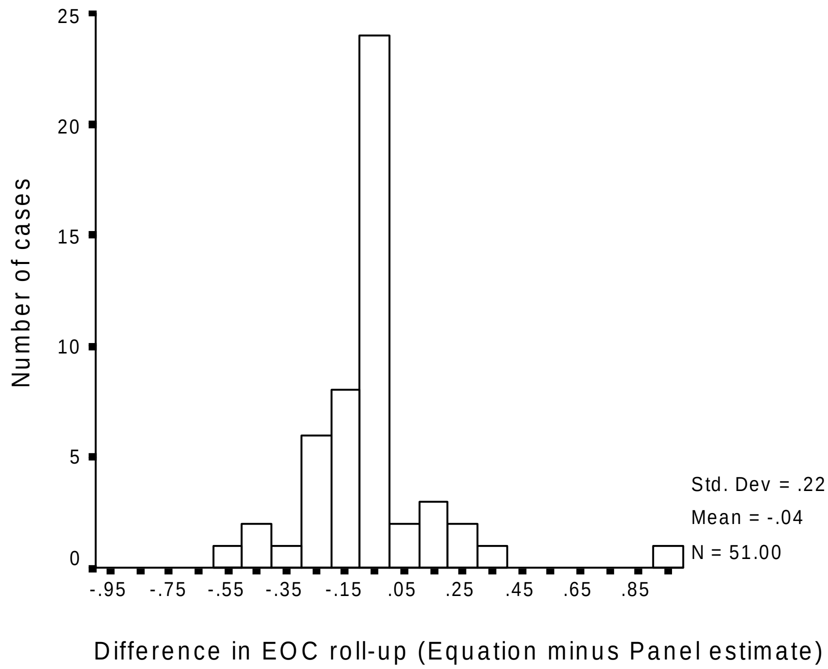


Figure 3. Difference in EOC code roll-up and panel estimate: panel criticalities and panel weights



An examination of figure 2 shows that the roll-up EOC code calculated by using the NWAS equation and NWAS parameters differs greatly from the roll-up EOC that the panel estimated. In general, the calculated EOC code is higher than the panel's estimate and in many instances, it is much higher. Note that the standard deviation of the difference is 0.40. This is enough to move a rating by two operational categories on the EOC metric—for example, from "major problem" to "minor problem."

In contrast, in figure 3 we see that the roll-up EOC code calculated by using the NWAS equation and the panel's parameters agrees reasonably well with the direct estimate made by the panel. In later sections, we will discuss ways that we think this agreement can be further improved.

We summarize relevant statistical parameters from the data sample in table 9.

Table 9. Summary statistics for the data sample

Criticality	Weight	Level	Cases	Difference between EOC from roll-up equation and the direct panel estimate: Mean	Standard error	Standard deviation
NWAS	NWAS	1	6	.39	.10	.24
NWAS	NWAS	2	36	.12	.07	.45
NWAS	NWAS	3	9	.06	.08	.24
		Total	51	.15	.06	.40
NWAS	Panel	1	6	.30	.12	.28
NWAS	Panel	2	36	.12	.05	.30
NWAS	Panel	3	9	.05	.12	.36
		Total	51	.13	.04	.31
Panel	Panel	1	6	-.02	.08	.18
Panel	Panel	2	36	-.07	.03	.16
Panel	Panel	3	9	.05	.12	.37
		Total	51	-.04	.03	.22

Table 9 gives additional insight into the data. It shows large differences in mean values (0.15) and large standard deviations (0.40) with the EOC code roll-up using the NWAS parameters. The largest differences are in equipment levels 1 and 2. It is in these levels that NWAS designated all systems as non-critical. Level 1 (mean difference of 0.39) is particularly crucial because it is the final stage in the roll-up.

An examination of the middle group of three rows shows that using the panel's weights with NWAS criticalities does not make much difference. We believe that this is due not to similarities between the panel's weights and the NWAS weights but rather to the structure of the roll-up equation.

Lastly, the roll-ups that used the NWAS equation with both the panel's criticalities and weights appear to be fairly satisfactory across all levels of equipment. The mean difference (-0.04) between the output of the roll-up equation and the panel's estimate is surely within any reasonable margin of uncertainty. We believe that the

large standard deviation (0.37) shown for level 3 can be reduced by introducing the concept of "conditional criticality" into the model. We will discuss this point in a later section.

Redundant systems

From our examination of the model, we feel that the model could be considerably improved if we could find a better way to deal with redundant systems. These systems are designed with redundancy to allow the ship to function even if some of the systems go down. However, in many cases if several of the redundant systems fail, the ship condition becomes severely degraded. Examples of such systems include propulsion, main propulsion 1, firemain, MK 84 400Hz SFC, and power plant.

Table 10 lists the model parameters for propulsion. NWAS designates all three subsystems (controls, main propulsion 1, and main propulsion 2) as non-critical. It is a separate issue as to whether main propulsion 1 and main propulsion 2 are individually critical, but, clearly, the combination is critical. With both units out of commission, the ship cannot move. Surely that condition would be critical to the EOC of the propulsion system.

Table 10. Level 2 redundant systems: propulsion

System	Criticality		Weight	
	NWAS	CNA panel	NWAS	CNA panel
Controls	N	C	30	34
Main Propulsion 1	N	N[a]	5	33
Main Propulsion 2	N	N[a]	10	33

a. These systems may be considered to be non-critical by themselves; however their combination is surely critical.

One approach to the problem would be to incorporate the concept of "conditional criticality." The criticality of a system would depend on the condition (i.e., EOC) of any redundant systems, i.e., are they really available to back it up? For example, if the EOC code of main

propulsion 1 was below some level, say 0.4, then the criticality of main propulsion 2 could be reset to "critical" and vice versa.

Consider the effect on the roll-up using the hypothetical subsystem EOC codes as shown in table 11. Hypothetical subsystem EOCs for level-2 roll-up: propulsion.

Table 11. Hypothetical subsystem EOCs for level-2 roll-up: propulsion

System	Hypothetical EOC conditions					
	A	B	C	D	E	F
Controls	1.0	1.0	1.0	0.5	0.5	0.5
Main Propulsion 1	0.5	0.0	0.0	1.0	0.5	0.5
Main Propulsion 2	1.0	1.0	0.5	1.0	1.0	0.5

On the basis of the hypothetical EOC codes in table 11, we assigned conditional criticalities as shown in table 12.

Table 12. Conditional criticalities for hypothetical subsystem EOCs for level-2 roll-up: propulsion

System	Hypothetical EOC conditions					
	A	B	C	D	E	F
Controls	C	C	C	C	C	C
Main Propulsion 1	N	N	N	N	N	N
Main Propulsion 2	N	C	C	N	N	N

We then used these conditional criticalities in the NWAS roll-up equation and show the results in table 13. The change to conditional criticalities is seen to reduce the equation roll-up from .51 to .34. This is much closer to the panel's estimate 0f .20 for condition "C".

We made revised roll-up calculations for five systems containing redundant subsystems. The results are summarized in table 14.

Table 13. Comparison of level-2 roll-up EOC codes: propulsion

Parameters		Hypothetical EOC conditions					
Criticality	Weights	A	B	C	D	E	F
NWAS	NWAS	.94	.89	.78	.67	.61	.50
NWAS	Panel	.84	.67	.51	.83	.67	.50
Panel	Panel	1.0	.67	.51	.50	.41	.34
Panel with conditional criticalities	Panel	1.0	.67	.34	.50	.41	.34
Panel estimate of roll-up		.78	.67	.20	.90	.70	.55

As shown in table 14, the modifications for redundant systems left the mean differences the same at -.04, decreased the standard error of the mean differences from .03 to .02, and greatly decreased the standard deviation of the differences from .22 to .13. This result is considerably better than the results without the modifications for redundant subsystems and much better than the unmodified NWAS equation with NWAS model parameters.

The bottom line is that the NWAS equation with the panel's model parameters and our proposed modifications to treat redundant subsystems rolls up to an overall EOC code that is in reasonably close agreement with what our panel of naval engineers thought the hypothetical situation dictated.

Overview on model parameters

The findings of this section may be summarized as follows:

- The NWAS model parameters (criticality and weights) appear to be unrealistic for many items.
- The NWAS roll-up equation, used with the NWAS model parameters, does not yield realistic results.
- The NWAS equation with the panel's model parameters and our proposed modifications to treat redundant subsystems does yield reasonably realistic results.

Table 14. Summary statistics for the data sample incorporating proposed changes in the treatment of redundant systems

				Difference between EOC from roll-up equation and the direct panel estimate		
Criticality	Weight	Level	Cases	Mean	Standard error	Standard deviation
NWAS	NWAS	1	6	.39	.10	.24
NWAS	NWAS	2	36	.12	.07	.45
NWAS	NWAS	3	9	.06	.08	.24
		Total	51	.15	.06	.40
NWAS	Panel	1	6	.30	.12	.28
NWAS	Panel	2	36	.12	.05	.30
NWAS	Panel	3	9	.05	.12	.36
		Total	51	.13	.04	.31
Panel	Panel	1	6	-.02	.08	.18
Panel	Panel	2	36	-.07	.03	.16
Panel	Panel	3	9	.05	.12	.37
		Total	51	-.04	.03	.22
Panel/ redundant systems	Panel	1	6	-.02	.08	.18
Panel/ redundant systems	Panel	2	36	-.06	.02	.13
Panel/ redundant systems	Panel	3	9	.02	.01	.04
		Total	51	-.04	.02	.13

We have evaluated the model using hypothetical EOC codes. In operational use, the model output will be only as good as the EOC codes assigned by technicians to individual low-level pieces of equipment. It would improve the credibility of the model output if the reliability of these assigned EOC codes were to be assessed and found to be high. We recommend that a selection of equipment be evaluated by a large number of technicians and that the method of repeated measurement of the same pieces be used to empirically measure the reliability of the evaluations.

Fundamental design and implementation issues

Consequences of arbitrary choice of scale

It is tempting to suppose that, in constructing an index like the EOC scale, there is some leeway in choice of structure and scale. For instance, it seems sensible to expect that, if a weighting scheme is used to combine different components of an index, a simple multiplicative rescaling of the weights shouldn't change any results. In the case of the EOC scale, this is true: If the weights are constrained to fall between (say) zero and one instead of zero and 100, the index will work exactly as it always did so long as the old weights are rescaled appropriately. Any configuration that was better than another under the old (0–100) scale would be better under the new one too.

However, there are limits as to what can be arbitrarily specified without affecting substantive conclusions. In this regard, we would draw particular attention to the EOC scale and to the gradations within it. This scale is seemingly innocuous and reasonable looking, but by imposing it, we are building in implicit rankings among state configurations that may or may not be valid. This is a subtle point, so to illustrate it in a concrete way, we now present a simple hypothetical example.

Imagine two identical systems, A and B, each having three subsystems of equal importance. (We will suppose equal weights for convenience and for the same reason will suppose that none of the subsystems are critical in the sense used in EOC roll-ups.) Suppose further that in System A, one of the subsystems is in perfect working order, a second is completely inoperable, and a third is at a level of operational capability just above inoperable. These states would correspond to EOC codes of 1, 0, and 0.3, respectively, under the current EOC scale described above. The rolled-up EOC code for System A would be the simple average of the three subsystem EOC codes or 0.43.

Now suppose that in System B all three of the subsystems are at EOC codes of 0.3, meaning that they are all barely functional. Thus, the EOC code for System B would also be 0.3—less than that of System A. Therefore, we would say that System A is in better condition than System B.

To see how this conclusion is driven by the choice of scale, let us consider an alternative EOC schedule as shown in table 15. The general idea of the scale in table 15 is that it is supposed to be similar to the current EOC scale except that the boundary between up and down occurs higher in the scale, at around 0.5, rather than at 0.2 as in the current scale. The descriptions shown are notional; it could be argued that we are mixing operational and maintenance perspectives in this case, but the labels attached to any particular intermediate value in the scale aren't really important for present purposes so long as it is understood that 0.6 is the lowest possible number for a functioning unit.

Table 15. Alternative EOC scale

Description	EOC
Perfect condition	1.0
Slightly impaired	0.9
Some problems	0.75
Barely operational	0.6
Borderline up/down	0.5
Down but easily fixed	0.4
Harder to repair	0.25
Major repair effort needed	0.1
Irreparable (replace)	0.0

When we apply the alternative scale of table 15 to Systems A and B from before, we find that System A has an EOC code of 0.53 and System B has an EOC code of 0.6. Thus, using the alternative scale, we would conclude that System B is in better condition than System A. This is shown in table 16.

Table 16. Comparing roll-up EOC for two different scales[a]

Subsystem	System A	System B
	Index 1:	
1	Perfect (1)	Barely operable (0.3)
2	Inoperable (0)	Barely operable (0.3)
3	Barely operable (0.3)	Barely operable (0.3)
Overall index	0.43	0.3
	Index 2:	
1	Perfect (1)	Barely operable (0.6)
2	Inoperable (0)	Barely operable (0.6)
3	Barely operable (0.6)	Barely operable (0.6)
Overall index	0.53	0.6

a. With a change in scales, the EOC for "barely operable" is changed from 0.3 to 0.6. With no other changes, the EOC of System B rises above the EOC of System A.

The point here is not to propose table 15 as an alternative to the existing EOC scale. It is rather to show that for this scale, its structure will affect the results of comparisons between systems (or ships, for example). Under the existing EOC scale, we would say that System A is in better condition than System B, whereas under the table 16 scale, we would conclude the opposite. Both of those conclusions cannot be correct; therefore, at least one of those scales must be wrong because in a simple case like this, we cannot blame incorrect weights or criticality determinations for the inconsistency. Thus, the question is, how do we know whether the current EOC scale (or indeed any other) is right?

As best we can determine, no one has examined this question; so far as our contacts at NAVSEA Corona were aware, there is no record of any analytical basis for the current EOC scale. We view this as a significant oversight in the development of this methodology, and we strongly recommend that it be given careful consideration. A few final comments on this subject:

- The previous discussion is essentially a theoretical argument. Although it is difficult for us to determine how important the choice of scale is in practice, we have been told that the FTSC technicians who do the inspections sometimes find it necessary to change the results of roll-up calculations because they give

answers that are obviously wrong. We can't help wondering whether these inconsistencies are due, at least in part, to scale problems. Also, if NWAS Corona isn't already doing so, it might be a good idea to keep systematic records of cases where changes are needed in order to better determine what the scale (and other structural characteristics) ought to be.

- One disturbing implication of this line of thinking is that there may not be one "right" scale; it is possible to envision situations where different scales may be necessary for different systems or ships.

- Thinking about these considerations brought us around to thinking about what we have come to call the "truth table" methodology; at least in principle, it provides a systematic method for thinking about appropriate scale and other structutal characteristics (such as functional forms). We turn to this approach in the next section.

Truth table methodology

One way to evaluate a set of metrics is to use a "truth table." For purposes of this paper, we define a truth table as a relationship among verbal variables in the same way that a function is a relationship among numerical variables. For example, if a car has a tire that is described as "flat" and brakes that "squeak and don't work well," the car might be described as "inoperable but readily reparable." A car with a bent frame might be described as "unreparable" or "totaled." The truth table is the relationship among the different descriptions of the car's parts and its whole.

A key question to ask about numerical measurements of material condition is whether the measurements are consistent with the truth table. For example, in the car example, the number assigned to "flat tire" and the number assigned to "brakes squeaking and don't work well" should result in an aggregate number that is consistent with "inoperable but readily reparable."

The truth table approach suggests further important questions in evaluating a system of metrics:

- *Is the dimensionality correct?* For example, inoperable but reparable has two dimensions: operability and fixability. Thus, a single number could not represent it.

- *Is the assignment of numbers to detailed problems consistent?* Are two conditions that result in the same numerical metric really of equal seriousness?

- *Is the aggregation correct?* For example, will one fatal problem correctly be evaluated as leading to an inoperable system?

We have not been able to apply this methodology fully under the current project. Nevertheless, we recommend that this methodology be applied, and there are some results from our initial investigation that suggest useful changes that could be made to the system of metrics.

When we examined the detailed assessment of individual parts of each system published as part of the evaluation for the USS *Lake Champlain,* we found that the evaluation listed a number of incommensurate deficiencies. Consider main propulsion as an example.

One particular main propulsion deficiency was listed as "flange shield improperly installed" on the piping for the fuel oil server system. The system impact is listed as "none," and the follow-up activity is listed as SFWL (Ship's Force Work List).

This example shows several dimensions:

- The extent to which the deficiency affects operations of the ship. Many would consider this dimension to be the only one that's appropriate for a measure of material condition.

- The extent to which the deficiency reflects other problems such as disorder and not following instructions.

- Who should take the follow-up action? Most of the deficiencies could be corrected by the ship's force. Others could be corrected by an intermediate activity (IMA). Few, if any, required shipyard intervention. Thus, care needs to be taken that this system not be used to evaluate how much shipyard work is needed and for what ships. Most of the information is not relevant for shipyard work.

Given the above findings, our recommendations are:

- *Consider disaggregating the overall index. We suggest a breakdown like the following*:
  - Operational problems that need shipyard work
  - Operational problems that need IMA work
  - Operational problems that can be solved by the ship's force
  - Problems that are not operational problems, such as improper installation, missing labels, and oil leaks.
- *Have the entire system of metrics reviewed from a truth-table point of view* to validate the mathematical formulas for computing the aggregate indexes. This would involve a simple listing of the problems of a ship or a ship's system. A panel of experts would then describe verbally their characterization of the ship or system as a whole. The question would then be whether the mathematics could replicate this relationship between the problems listed and the overall evaluation of the ship.

Connection with ERP

The discussion up to this point has taken the data-gathering methodology as given, at least in broad outline. We have not critically examined the basic idea of once-per-IDTC inspections conducted by a team visiting the ship. However, in light of planned improvements in NAVSEA's maintenance work information systems—specifically the NEMAIS Enterprise Resource Planning (ERP) project—we feel it makes sense to pause and consider whether the present method of data collection is still suitable given available information technologies and the planned upgrades.

The first phase of the NEMAIS project is expected to go live later this year and, as of this writing, work to extend the ERP system to the ships themselves has not yet begun in earnest. However, in informal conversations with NEMAIS personnel, we have been told that they expect to begin work on pilot instances of shipboard ERP configurations before the end of this year. When fully implemented, NEMAIS

will represent a real-time data stream about shipboard maintenance activity and requirements that has no counterpart in the Navy today. (The CASREP system has utility but does not provide the comprehensive visibility of maintenance activity that ERP will bring.)

In view of these expected advances, we believe that it would be wasteful and a duplication of effort to construct a system for measuring ship material condition that fails to take advantage of the data flows that the ERP system will produce as a matter of course. The existing system of infrequent inspections and reports could be replaced with a system in which shipwide roll-ups are available almost continuously and in something close to real time. We therefore strongly recommend that future SMCM development be done in coordination with NEMAIS.

Conclusions

Broadly speaking, our conclusions fall into two groups. The first of these takes the broad structure of the EOC code/rollup methodology as given and asks whether the parameters that have been developed and the results that have been obtained seem to be in line with what we would expect. The general conclusion here is that some of the specific parameters we looked at seemed out of line and that it would probably be a good idea to have them checked by independent sources such as shipyard engineers, port engineers, and engineering officers. Other validity and consistency checks, such as an experiment to determine the reliability of observed EOC measurements prepared by FTSC technicians, would be warranted, and we would also recommend that further attention be given to the question of how to handle redundant systems. However, none of these matters seems to us to be a fatal flaw in the system.

The other group of conclusions stems from our effort to "step back" and reconsider the SMCM approach from first principles. Here we find issues that are more far-reaching, and although we would not go so far as to conclude that the SMCM initiative has gone down an unproductive path, there are some fundamental issues that need to be addressed. These include the question of how to develop an appropriate EOC scale, whether one such scale is sufficient (as opposed to the necessity of developing multiple scales for different systems, for example), whether multi-dimensional metrics would be appropriate, and what those dimensions should be.[2]

2. Having directed our attention mainly to scale-related questions, we have largely ignored the matter of functional forms in roll-up equations. However, it seems likely that the same questions that we ask about arbitrarily-specified EOC scales could also be asked about those functions, and it would appear that the answers are the same: namely, that arbitrarily specified roll-up methods implicitly assume particular engineering relationships between systems that may or may not correspond to reality.

Finally, and in our view most importantly, we would point out that the ERP initiative currently being implemented by NAVSEA would provide an ideal vehicle for producing data flows concerning material condition of ships. It ought to be possible to use these to provide ship material condition metrics on a much more frequent and timely basis than the current approach, which is based on infrequent inspections. We therefore recommend that future development of the SMCM be done in conjunction with NAVSEA's NEMAIS project.

List of figures

List of tables

Distribution list
CRM D0005586.A2

N8

N81

N81B

N814

N4

N41

N43

N430

NAVSEASYSCOM

CINCLANTFLEET

CINCPACFLEET